



(12) **EUROPEAN PATENT APPLICATION**

(21) Application number : **93300962.3**

(51) Int. Cl.⁵ : **H04N 5/14, H04N 5/232**

(22) Date of filing : **10.02.93**

(30) Priority : **15.02.92 JP 61286/92**
17.02.92 JP 61123/92
25.02.92 JP 73320/92

 (43) Date of publication of application :
25.08.93 Bulletin 93/34

 (84) Designated Contracting States :
DE FR GB

 (71) Applicant : **SONY CORPORATION**
7-35, Kitashinagawa 6-chome Shinagawa-ku
Tokyo (JP)

(72) Inventor : **Kageyama, Koji, c/o Sony Corporation**
7-35 Kitashinagawa 6-chome
Shinagawa-Ku, Tokyo (JP)

 (74) Representative : **Nicholls, Michael John**
J.A. Kemp & Co., 14, South Square, Gray's Inn
London WC1R 5LX (GB)

(54) **Picture processing apparatus.**

(57) In a picture processing apparatus such as a television camera, monitor camera, television conference system and the like, an imaging device is controlled to follow desired target object and image it in the best state. In first target object following apparatus, a distribution of motion vectors within a tracking vector detection area is detected, and then the most characteristic motion vector is detected. In second target object following apparatus, a power of components in high frequency band is detected per an unit of predetermined blocks, as well as difference data between picture data of motion pictures of the present frame and the previous frame is obtained, the power of the difference data is detected per an unit of predetermined blocks, and an edge having movement is detected on the basis of the detection result. In a picture conversion apparatus by which high quality motion picture is transmitted, such as a picture processing apparatus in which picture data is processed to motion compensate a picture data, a motion field detection circuit detects disorder of a motion vector, on the basis of which a quantizing step size is changed so that the step size is roughly set regarding to a portion of picture not to easy perceive a deterioration of a picture quality.

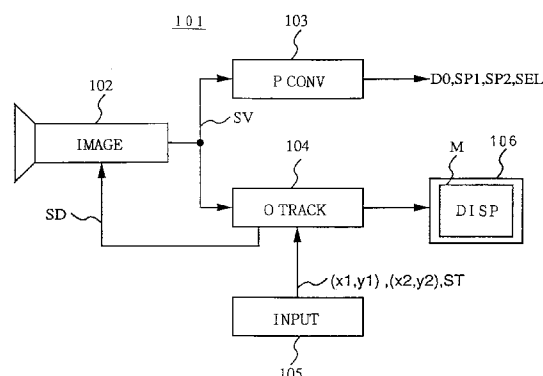


FIG. 1

This invention relates to a picture processing apparatus, and is applicable to an imaging device, for example, a television camera, a monitor camera, and a television conference system, besides a picture data transmitting system, which carries out compensation for the motion of picture data, and the like.

Conventionally, as an apparatus for following an object to be imaged, was proposed a picture processing apparatus for following a targeted portion of a picture, by detecting the brightest region in an image picked up, on the basis of which desired target portion is followed.

If any desired portion of an image can be tracked in such a way, then automatic focusing, iris adjustment and the like can be performed for this portion of the image to be able to achieve the improvement the easy use of imaging devices.

However, there is a problem that the apparatus in which a desired portion of an image is tracked based on the brightest region in the image, can not follow a target portion of the image reliably, thereby being insufficient to use practically.

Method for solving this problem is to follow the target portion of an image on the basis of picture characteristic thereof. However, this method has a disadvantage that a generally complicated structure is required for the apparatus.

Further method for solving this problem is to detect motion of target portion of an image, by detecting a power of difference data obtained from difference between frames, but in some cases, a detected result which is different from the motion of target portion of an image may be obtained.

It is proposed other method which is to detect motion of target portion of a picture, by using a time spatial filter apparatus. However, in this method, amount of operation is to become large, therefore, this method has a disadvantage that generally complicated structure is required for the apparatus.

On the other hand, conventionally, a method for transmitting a motion picture with high efficiency is proposed.

More specifically, in such a picture conversion apparatus, picture data sequentially inputted are divided into predetermined unit blocks, and then a motion vector is detected for every block, by using the block matching method.

Then, residual data is generated by subtracting a picture data sequentially motion compensated each picture data, after motion compensating a picture data which is for one frame, on the basis of said detected result.

Thus, data amount of picture data is previously to be reduced, in comparing with a case which picture data is directly transmitted by using correlation of between frames of motion picture.

This residual data is sequentially processed with discrete cosine transform and variable length coding, and is stored into a buffer memory after re-quantizing said residual data.

Thus, the picture data stored in the buffer memory are sequentially outputted from it at predetermined timing, and picture data is outputted with keeping constant value amount of data, then by changing a quantization step size of re-quantizer, corresponding to a residual quantity of said buffer memory.

Thus, the motion picture can be transmitted with high efficiency, by using correlation of between frames of motion picture.

Now, in the motion picture, a part of large motion has a characteristic that deterioration of picture quality is imperceptible.

Further, in such motion picture, having a case where part of a picture pronounced changes compared with changes of thus around picture, another case of changes whole of a picture.

In this case of motion picture, it has a characteristic that deterioration of picture quality is more imperceptible, in comparing with a case where whole of a picture has constant motion.

With effectively using this unique characteristic, it is expected that more efficient transmission of motion pictures can be achieved. It is also expected that the quality of motion picture can be improved compared to the conventional technology.

In view of the foregoing, an object of this invention is to provide a picture processing apparatus, for example, television camera, monitor camera, television conference system and the like, which controls an imaging device to take a picture into the best condition with tracking a target portion of a picture. Further, in thus obtained picture, high quality motion picture is to be transmitted with high efficiency, by making effective use of human visual sensitivity.

A target object following apparatus for following a portion of a picture, comprises: an input circuit for setting a tracking vector detection region to a motion picture inputted from an imaging device;

a tracking vector detection circuit for detecting a motion direction and motion distance of a picture in the tracking vector detection region, about a picture in the tracking vector detection region which is set by said input circuit;

a window-frame data generation circuit for designating an area followed by said tracking vector detection circuit by use of window frame, and for generating a control signal which always to be caught a target portion

of a picture;

a picture composing circuit for composing a motion picture inputted from said imaging device with said window frame; and

5 a camera driving circuit for driving said imaging device, by controlling signal generated from said window-frame data generation circuit; wherein

The output of the picture composing circuit is displayed on a picture displaying device to be able to observe a state of following the target object, as well as controls the imaging device to be able to catch the target object always in the best state.

10 The tracking vector detection circuit can track the target portion of a picture, by detecting a distribution of motion vector in said tracking vector detection region, and can track target portion of a picture by using a spatial edges and a time edges.

On the other hand, a picture conversion apparatus for performing transmission of said high quality motion picture, comprises:

15 a motion detection circuit for detecting a motion vector each blocks, after dividing a picture data inputted from an imaging device at predetermined unit blocks;

a picture data processing apparatus for outputting a residual data with motion compensating a picture data on the basis of a motion vector;

a motion field detection circuit for detecting disorder of motion vector about a block of residual data and adjacent blocks in surroundings;

20 a re-quantization circuit for outputting a quantization data with quantizing the residual data; and

a buffer circuit for with storing a quantization data and outputting thereof; wherein

when motion vector is disorder, the re-quantization circuit quantizes the residual data, after changing the quantization step size in a large way, on the basis of the detection result from the motion field detection circuit.

25 With the above construction of the present invention, it can be detected a movement direction and movement distance of target portion of a picture with simple construction, and can be achieved by making effective use of human visual sensitivity so as to transmit the picture signal with high picture quality.

The invention will be further described by way of example in the following detailed description when read in conjunction with the accompanying drawings in which like parts are designated by like reference numerals or characters, and in which:-

Fig. 1 is a block diagram showing a picture processing apparatus according to the present invention;

Fig. 2 is a block diagram showing first and second embodiments of a target object following apparatus;

Fig. 3 is a schematic diagram illustrating description of setting of a tracking vector detection region, in the first and second embodiments of the target object tracking apparatus;

35 Fig. 4 is a schematic diagram illustrating description of a frame display position, in the first and second embodiments of the target object tracking apparatus;

Fig. 5 is a block diagram showing a tracking vector detection circuit, in the first and second embodiments of the target object tracking apparatus;

40 Fig. 6 is a block diagram showing a motion detection circuit, in the first embodiment of the target object tracking apparatus;

Fig. 7 is a schematic diagram illustrating description of a blocking circuit, in the first and second embodiments of the target object tracking apparatus;

Fig. 8 is a schematic diagram illustrating a description of detecting of motion vector, in the first embodiment of the target object tracking apparatus;

45 Fig. 9 is a block diagram showing a tracking vector detection circuit, in the second embodiment of the target object tracking apparatus;

Fig. 10 is a block diagram showing an embodiment of a picture conversion apparatus;

Fig. 11 is a schematic diagram illustrating a block of picture data, in the embodiment of the picture conversion apparatus;

50 Fig. 12 is a schematic diagram illustrating a motion vector detection region, in the embodiment of the picture conversion apparatus;

Fig. 13 is a schematic diagram illustrating description of motion field detection, in the embodiment of the picture conversion apparatus;

Fig. 14 is a schematic diagram illustrating a case when motion vectors are small disorder;

55 Fig. 15 is a schematic diagram illustrating a case when motion vectors are large disorder;

Fig. 16 is a schematic diagram illustrating description of switching a quantization step size, in the embodiment of picture conversion apparatus; and

Fig. 17 is a block diagram showing a decoder, in the embodiment of the picture conversion apparatus.

Preferred embodiments of this invention will be described with reference to the accompanying drawings:

(1) General Construction of Picture Processing Apparatus In The Present Invention

5 In Fig. 1, 101 designates a general construction of a picture processing apparatus in the present invention. Video signal SV imaged by an imaging device 102 is transmitted to a picture conversion apparatus 103 and an target object tracking apparatus 104.

An input circuit 105 sets a position of a target portion of a picture by use of frame, and outputs coordinates data shown thus tracking vector detection region (x1, y1) (x2, y2) and trigger signal ST which commands track-
10 ing execution.

In the target object tracking apparatus 104, a target portion of a picture selected by the input circuit 105 is tracked, as thus tracking state, a picture having a shape of a window frame which catches the target portion of a picture is composed with video signal SV, then thus composite picture M is displayed on a picture display 106.

15 In this way, the video signal SV imaged into the best condition by imaging device 102, is performed to effectively picture signal process by use of human visual characteristic in a picture conversion apparatus 103, thereby output picture signals D0, SP1, SP2 and SEL are transmitted.

(2) First Embodiment of the Object Tracking Apparatus

20

(2-1) General Construction

In Fig. 2, 201 designates the target object tracking device for tracking a picture of a subject specified by a user to adjust an iris and focus.

25 That is, an imaging device 102 images a picture of a desired target object in a field of vision and then outputs a resultant video signal SV of the picture to a tracking vector detection circuit 202.

The tracking vector detection circuit 202 converts successively input video signal SV into a digital signal and sets a tracking vector detection region AR, based on coordinate data (x1, y1) and (x2, Y2) input from an input circuit 105.

30 Furthermore, when a trigger signal ST is input from the input circuit 105, the tracking vector detection circuit 202 detects a tracking vector VT which represents a motion direction and motion distance of the target object of a portion of a picture existing in the tracking vector detection region AR, and furthermore outputs the detection result to a window-frame data generator 205.

35 When the tracking vector VT is obtained, the tracking vector detection circuit 202 shifts the tracking vector detection region AR by the amount corresponding to the tracking vector VT. Then, the tracking vector detection circuit 202 repeats the detection of the tracking vector VT for the subsequent frame using the tracking vector detection region AR which has been shifted.

40 In this way, tracking-vector detection circuit 202 detects a motion direction and motion distance of a picture in the tracking vector detection region AR and also shifts the tracking vector detection region AR to follow the motion of the picture.

In the target object tracking device 201, the input circuit 105 is operated for inputting coordinate data with a joystick so that the tracking-vector detection region AR is set by operation with the joystick.

45 In the target object tracking device 201, the video signal SV is applied to a picture composing circuit 204, whereas a picture having a shape of a window frame (hereafter referred to as a window-frame picture) WK is generated based on a window-frame data DW given from the window-frame data generator 205 as shown in Fig. 3.

Moreover, in the target object tracking device 201, the window-frame picture WK is superimposed on the video signal SV by the picture composing circuit 204, then the resultant composite picture M is displayed on a picture display 106.

50 In the above process, based on the coordinate data (x1, y1) and (x2, y2) input from the input circuit 105, the window-frame data generator 205 generates a window-frame data DW corresponding to the tracking-vector detection region AR so that the window-frame picture WK represents the tracking-vector detection region AR.

In this way, in the target object tracking device 201, a tracking-vector detection region AR can be set easily only by operating a joystick with monitoring a picture display 106.

55 In this preferred embodiment, a window-frame picture WK of predetermined size is displayed at the central position of the screen in the initial state. Then, a user can change the position and size of the window-frame picture WK by operating the joystick and monitoring the displayed picture to set desired tracking-vector detection region AR.

In this setting operation, the tracking-vector detection region AR is set to a size including an certain integer number of unit blocks for detecting a motion vector so that a tracking vector can be detected easily based on the detection result of the motion vector.

Moreover when a starting button for starting the tracking operation is pressed, the input circuit 105 outputs a trigger signal ST for starting the tracking operation, whereas an ending button for ending the tracking operation is pressed, the input circuit 106 outputs a control signal for ending the tracking operation. In this way, operation is controlled overall for the targeted-object tracking apparatus 1.

As shown in Fig. 4, when a tracking vector VT (DX, DY) is obtained, the window-frame data generator 205 adds two-dimensional coordinate components DX, DY to the coordinates (x1, y1) and (x2, y2) of the window-frame data DW and outputs the result.

Thus, based on the tracking vector VT, the window-frame data generator 205 updates the window-frame data DW so that the window-frame picture WK is moved to follow the motion of an picture in tracking-vector detection region.

Furthermore, in this embodiment, the window-frame data generator 205 outputs the window-frame data DW to a control circuit (not shown), then this control circuit detects a signal level of the video signal SV in the region which is defined by the window-frame data DW (that is, the tracking-vector detection region AR). Based on the detection result, the control circuit controls the focus and the iris of the imaging device 102.

In this way, in the target object tracking apparatus 201, the motion of the targeted picture in the tracking-vector detection region AR which is set by a user is tracked so that the optimum brightness and focus are always maintained for the targeted picture.

(2-2) Tracking-Vector Detection Circuit

As shown in Fig. 5, the tracking-vector detection circuit 202 gives the video signal SV to the motion detection circuit 501, by which the motion vector MV is detected.

Then, as shown in Fig. 6, the motion vector detection circuit 501 gives the video signal SV to the blocking circuit 601. The motion vector detection circuit 501 converts this video signal SV into a digital signal (hereafter, referred to as a digital video signal).

Furthermore, as shown in Fig. 7, the blocking circuit 601 divides the digital video signal into blocks BL each of which contains 8 x 8 pixels, then outputs the digital video signal for every block according to the order of raster scanning (Fig. 7 shows a block located at (m, k) pixel position).

A matching circuit 602 obtains the motion vector MV for every block by adapting the block matching method to the result of comparison between the output signal from the blocking circuit 601 and the output signal from a frame memory 603.

That is, as shown in Fig. 8, the matching circuit 602 assigns a region which is wider than a block by 2Dx x 2Dy in horizontal and vertical directions respectively for a motion-vector detection region ARM for every block. Then, the matching circuit 602 detects the difference between the present picture data in a block and the one-frame-previous data in the motion-vector detection region ARM, wherein the one-frame-previous data is stored in the frame memory 603.

During the above process, the matching circuit 602 changes successively the location of the block BL within the motion-vector detection region ARM, and determines the location where the sum of the absolute values of the above differences is minimized.

In this way, the matching circuit 602 detects the motion vector MV for every block BL, and then outputs the detected result to a distribution detecting circuit 502.

The distribution detecting circuit 502 detects the distribution of the motion vector MV in the tracking-vector detection region AR with respect to the direction and magnitude. The detected result is output to a judging circuit 503.

More specifically, the distribution detecting circuit 502 has a table having a size of 2DX x 2DY corresponding to the motion vector detection region ARM. The contents of the table are initialized and set to 0 at the beginning of every frame. The distribution detecting circuit 502 updates the contents of the table by one at a time depending on the value of the motion vector MV. Thus, the distribution of the motion vector MV is detected based on the motion-vector detection region ARM.

Furthermore, the distribution detecting circuit 502 updates the tracking-vector detection region AR, based on the detection result with respect to the tracking vector given by the judging circuit 503. Thus, while the distribution detecting circuit 502 updates the tracking-vector detection region AR for every frame to track the motion of the picture, it detects the distribution of the motion vector in the update tracking-vector detection region AR.

Based on the result of the detected distribution, the judging circuit 503 sets a motion vector MV having

the biggest count in the distribution as the tracking vector VT, then outputs this tracking vector.

A motion vector MV showing the most dominant motion among the motion vector within a frame which is set to surround the target object can be considered to represent the direction and magnitude of the motion of the target object.

5 Therefore, by detecting the most dominant motion vector, the desired target object can be tracked with a simple construction of apparatus. This tracking capability can be effectively used for automatic focusing and automatic adjustment of an iris. Thus, the easier operation is achieved for the target object tracking apparatus 1.

10 (2-3) Operation of an Embodiment

In the above construction, the output video signal SV from the imaging device 102 is applied to the tracking-vector detection circuit 202. Then, after converting the video signal into the digital signal, the blocking circuit 601 divides the digital video signal into blocks consisting of 8 x 8 pixels.

15 For each of these blocks, the motion vector MV is detected by the matching circuit 602.

This motion vector MV is applied to the distribution detecting circuit 502, by which the distribution of the motion vector MV in the tracking-vector detection region AR is detected and a motion vector MV showing the most dominant motion is set as the tracking vector VT.

20 As described above, the distribution of the motion vector is obtained and the most dominant motion vector MV is set as the tracking vector VT, thus the motion of the target object can be tracked with a simple construction.

The detection result of the tracking vector is applied to the window-frame data generator 205, by which the window-frame data DW is updated based on the tracking vector VT so that the window-frame picture WK is moved on the screen of the display 106 following the motion of the picture.

25 Moreover, by detecting the signal level of the video signal SV in the region defined by the window-frame data DW, it is possible to control the focus and the iris of the imaging device 102, following the motion of the target object.

Furthermore, the detection result of the tracking vector is applied to the distribution detecting circuit 502, by which the tracking vector is detected for the subsequent frame. In this way, the tracking vector can be detected for successive frame.

(2-4) Other Embodiments

35 In the above embodiment, the focusing and the iris are adjusted based on the detection result of the tracking vector. However, the present invention is not limited to that. Zooming operation or the like can be also performed in addition to that.

Furthermore, instead of moving the window frame, similar operation can be achieved by outputting the control signal SC to the camera driving circuit 206 (Fig. 2) to change the picture sensing region following the motion of the target object. In this case, combination of the motion of the window frame and change of the picture sensing region is also possible to adapt to the motion velocity.

40 In the above embodiment, the contents of the table are updated directly from the detection result of motion vector to detect the tracking vector. However, the present invention is not limited to that. After re-quantization of the motion vector, similar operation can be also achieved.

45 In this case, when the accuracy of the tracking vector is not good enough, the size of the table can be reduced. Thus, it becomes possible to detect the motion of the target object with a simpler construction.

Furthermore, in the above embodiment, after detecting the motion vector MV for every block BL, the distribution of the motion vector MV in the tracking-vector detection region AR is detected. However, the present invention is not limited to that. Similar effect can be also achieved for the case where the distribution is detected after the motion vector MV is detected only within a block in the tracking-vector detection region.

(3) Second Embodiment of The Target Object Tracking Device

(3-1) General Construction

55 Second embodiment of the targeted-object tracking apparatus, as described above in Fig. 2, has whole of construction same as first embodiment of said target object tracking apparatus of that.

(3-2) Tracking-Vector Detection Circuit

The tracking-vector detection circuit 202 detects motion edges from the video signal SV, then based on this detection result detects the tracking vector VT.

That is, as shown in Fig. 9, in the tracking-vector detection circuit 202, the successively input video signal SV is converted into a digital signal to generate an picture data, and then this picture data is applied to a blocking circuit 901.

The blocking circuit 901 successively outputs the above picture data, according to the order of raster scanning, together with the picture data with regard to eight pixels surrounding the present picture data.

In this way, the blocking circuit 901 divides outputs the successively input video signal SV into unit blocks containing 3 x 3 pixels, then successively outputs the video signal block by block.

A deviation detecting circuit 902 determines the average value with respect to the output data from the blocking circuit 901, then generates the deviation data by subtracting each output data from the above obtained average value.

Thus, the deviation detecting circuit extracts high frequency components block by block and outputs the results.

Furthermore, the deviation detecting circuit 902 determines the sum of the absolute values of the deviation data to determined the sum of the absolute values of the deviations for each block. Thus, the high frequency component power is determined block by block.

A spatial edge detecting circuit 903 compares the power detected by the deviation detecting circuit 902 with a predetermined threshold stored in a table 904 and set a spatial edge detection flag KF when the high frequency component power becomes larger than the threshold.

That is, when there exist edges, a large value of the deviation data is obtained for each corresponding block.

Therefore, as in this specific embodiment, if the high frequency component power is detected for each block, then it becomes possible to detect the existence of an edge, based on the power detection result.

In this way, in the tracking-vector detection circuit 202, the high frequency component power is detected for predetermined blocks to detect edges in the two-dimensional space of a motion picture.

A frame memory 905 stores and outputs one frame of digitized video signal SV. Thus, the frame memory 905 can output an picture data which is older by one-frame than the output data of the blocking circuit 901 (hereafter, the data in the frame memory will be referred to as a previous frame data and the output data of the blocking circuit 901 will be referred to as a current frame data).

In the above processing, the frame memory 905 successively outputs the picture data corresponding to the blocking circuit 901, together with the data of eight pixels surrounding the current pixel, according to the order of the raster scanning.

Thus, the frame memory 905 successively outputs the previous frame data, corresponding to the output data of the blocking circuit 901, block by block wherein one block consists of 3 x 3 pixels.

A difference detecting circuit 906 obtains the difference data by subtracting the output data of the frame memory 905 from the corresponding output data of the blocking circuit 901. Thus, the difference detecting circuit 906 successively outputs the difference data for each 3 x 3 pixel block.

A power detecting circuit 907 detects the power of the difference data by detecting the sum of the absolute values of the difference data, and then outputs the obtained result.

A time edge detecting circuit 908 compares the power detected by the power detecting circuit 907 with a predetermined threshold stored in a table 909 and sets a time edge detection flag JF when the power of the difference data exceeds the threshold.

That is, when change occurs in a picture, a large value is obtained in the difference data between the current frame and the previous frame.

Obtaining the difference data between the current frame and the previous frame implies extracting of high frequency components along the time axis. This means that edge components are extracted with respect to the time axis.

That is, if the power of the difference data is detected for each small block consisting of 3 x 3 pixels, then, based on this detection result, it is possible to detect edges along the time axis, that is, an abrupt change in a picture as in the case where an edge moves in the two-dimensional space.

Based on the edge detection principle described above, an edge judging circuit 910 sets a motion edge detection flag JKF when the spatial edge detection flag KF and the time edge detection flag JF are set. Thus, the edge judging circuit 910 detects a motion edge and outputs the detected result.

In this way, edges are detected with respect to both of time and space using a small block as a detection unit. Based on this detection result, motion edges are detected. Thus, it is possible to easily make reliable de-

tection of motion edges.

As shown in Fig. 7, a motion detecting circuit 911 divides a digital video signal into blocks BL wherein each of block consists of 8×8 pixels (Fig. 7 shows a block located at the position $(k \times m)$), and then detects the motion vector a MV for each block by using the block matching method.

That is, the motion detecting circuit 911 detects the residual difference in the data between the previous frame and the current frame for each block, then further determines the sum of the absolute values of the above resultant residual difference data.

In this processing, the motion detecting circuit 911 successively shifts the position of each block within a predetermined motion vector detection region to detect the position where the above sum of the residual difference data is minimized. Thus, the motion vector is detected.

In the above processing, the motion detection circuit 911 selectively performs the above motion vector detection processing only for blocks whose corresponding motion-edge detection flag JKF is set in the tracking-vector detection region AR.

The selective detection of the motion vector only for the blocks in the tracking-vector detection region AR leads to reduction in the amount of calculation compared to the case where the motion vector is detected for all blocks.

Furthermore, by detecting the motion vector only for the blocks whose corresponding motion-edge detection flag JFK is set, it is possible to selectively detect only the motion vectors reflecting the motion of the target object. Thus, the detection accuracy can be improved. Besides, the amount of calculation can be also reduced.

A judging circuit 912 detects the distribution of the motion vector, based on the direction and magnitude of the motion vector. Then, the decision circuit 912 detects the tracking vector VT, based on the detection result of the above distribution.

That is, the decision circuit 912 has a table having a size of $2D_x \times 2D_y$ corresponding to the motion-vector detection region. The contents of this table are initialized and set to "0" at the beginning of every frame. The decision circuit 912 adds unit value, "1", at a time, to the contents of the table depending on the value of the motion vector MV. Thus, the distribution of the motion vector is detected with respect to the motion-vector detection region.

Based on the result of the detected distribution, the decision circuit 912 sets a motion vector MV having the biggest count in the distribution as the tracking vector VT, and then outputs this tracking vector.

A motion vector MV showing the most dominant motion among the motion vector within a window frame which is set to surround the target object can be considered to represent the direction and magnitude of the motion of the target object.

Therefore, by detecting the most dominant motion vector, the desired target picture can be tracked with a simple construction of apparatus. This tracking capability can be effectively used for automatic focusing and automatic adjustment of an iris. Thus, the easier operation is achieved for the target object tracking apparatus.

Furthermore, the judging circuit 912 updates the tracking-vector detection region AR, based on the result of the tracking vector detection. Thus, while updating the tracking-vector detection region AR for every frame to follow the motion of the picture, the judging circuit 912 detects the distribution of the motion vector in the updated tracking-vector detection region AR.

(3-3) Operation of an Embodiment

In the above construction, the output video signal SV from the imaging device 102 is successively applied to the deviation detecting circuit 902 by unit of 3×3 pixels via the blocking circuit 901. The deviation detecting circuit 902 determines the deviation of the given data from the average value.

Furthermore, the sum of the absolute values of the above deviation data is determined. Thus, the high frequency component power is successively detected in the order of raster scanning for the current pixel with respect to the eight pixels surrounding the current pixel. In the spatial edge detecting circuit 903, the above detected power is compared with a predetermined threshold and spatial edge detection flag KF is set when an edge is observed in the two-dimensional space.

Furthermore, the difference detecting circuit 906 determines the difference between the output data of the blocking circuit 901 and the picture data of the previous frame. Then, the power detecting circuit 907 detects the power of the above difference data.

In the time edge detecting circuit 908, this resultant power is compared with the predetermined threshold value. When there is observed an edge along the time axis, that is, an abrupt change in an picture as in the case where an picture moves on the two-dimensional space, the time edge detection flag JF is set.

The spatial edge detection flag KF and the time edge detection flag JF are applied to the edge judging

circuit 910. The edge judging circuit 910 sets the motion-edge detection flag JKF when both flags KF and JF are sets. Thus, the motion edge can be detected.

The motion edge detection flag JKF is applied to the motion detecting circuit 911, where the motion vector is detected based on the motion-edge detection flag JKF.

5 That is, in the motion detecting circuit 911, the video signal is divided into blocks each of which consists of 8 x 8 pixels, and then the motion vector is detected for the blocks whose corresponding motion-edge detection flag JKF is set in the tracking-vector detection region.

The motion vector MV is applied to the decision circuit 912. The decision circuit 912 detects the distribution of the motion vector and sets a motion vector having the largest count in the distribution as a tracking vector VT.

This tracking vector VT is applied to the window-frame data generator 205. Then, the window-frame data generator 205 updates the window-frame data according to the tracking vector VT so that the window-frame picture WK is moved on the screen of the display 106 following the motion of the target object.

Furthermore, by detecting the signal level of the video signal in the region defined by the window-frame data DW, it is possible to control the focus and the iris of the imaging device 102 following the motion of the target object.

(3-4) Other Embodiments

20 In the above embodiment, the power is detected for a region consisting of 3 x 3 pixels to detect the edge. However, the present invention is not limited to that. The detection region can be set in a broad range as required.

Furthermore, in the above embodiment, in addition to the picture data which is output according to the order of raster scanning, the picture data of eight pixels surrounding the above picture data is also output so that the picture data processing is successively carried out for each unit data consisting of 3 x 3 pixels. However, the present invention is not limited to that. For example, it is also possible to divide the motion picture into units consisting of 3 x 3 pixels for the processing.

In the above embodiment, the power of the difference data is detected by determining the sum of the absolute values of the difference data. However, the present invention is not limited to that. As in the case of the deviation data, the power of the difference data can be also obtained by calculating the sum of the absolute values of the deviation from the average value.

Furthermore, in the above embodiment, the power of each data is determined from the sum of the absolute values. However, the present invention is not limited to that. The high frequency component power can be also determined from the sum of squares.

35 In the above embodiment, the focusing and the iris are adjusted based on the detection result of the tracing vector. However, the present invention is not limited to that. Zooming operation or the like can be also performed in addition to that.

Furthermore, instead of motion the window frame, similar operation can be achieved by outputting the control signal SC to the camera driving circuit 206 (Fig. 2) to change the picture sensing region following the motion of the target object. In this case, combination of the motion of the window frame and change of the picture sensing region is also possible to adapt to the motion velocity.

In the above embodiment, to detect the tracking vector, the contents of the table are updated directly from the detected result of the motion vector. However, the present invention is not limited to that. A similar effect can be obtained by re-quantizing the motion vector.

45 In this case, when the accuracy of the motion vector is not good enough, the size of the table can be reduced. Thus, the detection of the motion of the target object can be achieved by a simpler construction.

Furthermore, in the above embodiment, the motion vector is detected only for the blocks whose corresponding motion-edge detection flag is set in the tracking-vector detection region. However, the present invention is not limited to that. The motion vector can be also determined by selectively detecting the distribution after detecting the motion vector for all blocks.

Alternatively, in the above processing, the motion-edge detection flag may be counted for every block and the distribution of the motion vector may be detected only for the blocks having a count larger than a predetermined number.

55 Furthermore, in the above embodiment, the target object is tracked by detecting the tracking vector. However, the present invention is not limited to that. The present invention can be preferably applied to various picture processing systems in which a motion picture is processed by detecting the motion edge.

(4) An Embodiment of The Picture Conversion Apparatus

(4-1) General Construction

In Fig. 10, 1001 generally designates a picture conversion apparatus which performs data compression for a successively input video signal, and then transmits the compressed video signal.

More specifically, in the picture conversion apparatus 1001, after a video signal is converted into a digital signal, this digital video signal is divided into predetermined unit blocks BL each of which has 8 x 8 pixels as shown in Fig. 11.

Furthermore, in the picture conversion apparatus 1001, these blocks BL are successively applied to a motion detection circuit 1003, by which a motion vector V is detected for each block BL using the block matching method.

That is, in the picture conversion apparatus 1001, one-frame-previous video data is stored in a frame memory 1004.

As shown in Fig. 12, the frame memory 1004 sets a region consisting of 40 x 24 pixels around the block BL which is input to the motion detection circuit 1003 as a motion vector detection area ARV. Then, the frame memory 1004 successively outputs the video data for every block BL in the motion vector detection area ARV.

The motion detection circuit 1003 detects the difference between the video data of the block BL and the output data from the frame memory 1004 so as to obtain the difference data between the video data of the block BL and the one-frame-previous video data, and then furthermore determines the sum of the absolute values of these difference data.

In the above processing, in order to obtain the motion vector (in this case, the motion vector consists of the values -X and -Y with respect to the horizontal and vertical directions, respectively) the motion detection circuit 1003 successively moves the block BL during the motion vector detection area ARV so as to determine the position A at which the sum of the absolute values of the difference data is minimized.

At the same time, the motion detection circuit 1003 also detects the sum of the absolute values of the video data for each block BL. Then, based on the comparison result between the above sum of the absolute values and the detection result of the motion vector, the motion detection circuit 1003 outputs a selection signal SEL for selecting an intraframe coding processing or interframe coding processing.

Thus, in the picture conversion apparatus 1001, depending on the amount of the data of each block BL, either the intraframe coding processing or the interframe coding processing is selected, and then the video signal of each block is transmitted. In this way, the video signal is transmitted efficiently.

That is, in the case where the direct transmission of the video data of each block BL requires a less amount of the data in transmission, the motion detection circuit 1003 selects the intraframe coding processing.

In this case, the motion detection circuit 1003 switches the connection of a selection circuit 1005 so as to output the video signal of each block BL directly to a re-quantization circuit 1006.

Furthermore, when the intraframe coding processing is selected, the motion detection circuit 1003 turns off a switching circuit 1007 so that the output data of the inverse re-quantization circuit 1008 is directly stored into the frame memory 1004.

After performing the discrete cosine transform for the output data from the selection circuit 1005, the re-quantization circuit 1006 re-quantizes the converted data, and then outputs the result of the re-quantization.

A variable length coding circuit (VLC) 1010 performs variable length coding for the output data from the re-quantization circuit 1006, then outputs the coded result to a buffer circuit 1011 so as to successively output the transmission data D0 at a predetermined transmission rate corresponding to the capacity of the transmission line.

In contrast, the inverse re-quantization circuit 1008 inputs the output data from the re-quantization circuit 1006, then performs the inverse processing to the processing performed by the re-quantization circuit 1006.

Thus, when the intraframe coding processing is selected, the inverse re-quantization circuit 1008 reproduces the input data given to the re-quantization circuit 1006, then stores the reproduced data into the frame memory 1004. In this way, one-frame-previous video signal is stored in the frame memory 1004.

On the other hand, in the case where motion compensation gives a less amount of data required for transmission, the motion detection circuit 1003 selects the interframe coding processing.

In this case, the motion detection circuit 1003 switches the connection of the selection circuit 1005 so that instead of the video data of each block BL, the output data from a subtracter 1012 is applied to the re-quantization circuit 1006.

Furthermore, when the interframe coding processing is selected, the motion detection circuit 1003 turns on the switching circuit 1007 so that the output data of a motion compensation circuit 1013 is added to the output data of the inverse re-quantization circuit 1008, and then this resultant addition is stored in the frame

memory 1004.

That is, based on the detection result of the motion vector V, the motion compensation circuit 1013 cuts an area containing 8 x 8 pixels (the area with hatched lines in Fig. 12) out of the motion vector detection area ARV, and then outputs the resultant video data DM of the cut area. Thus, the motion compensated video data is output.

The subtracter 1012 subtracts the motion compensated video data DM from the video data of each block BL, then outputs the motion compensated residual difference data to the selection circuit 1005.

In this way, in the picture conversion apparatus 1001, after the discrete cosine transform is carried out for the above residual difference data, the re-quantization is also carried out, and then the result is transmitted.

As described above, in the picture conversion apparatus 1001, by switching the intraframe coding processing and the interframe coding processing, efficient transmission of the video data D0 is achieved using the adaptive coding method.

Furthermore, when the interframe coding processing is selected, the switching circuit 1007 is turned on so that the motion compensated video data DM is applied to the adder 1009.

The adder 1009 adds the output data of the inverse re-quantization circuit 1008 with the motion compensated video data DM, thus reproduces the output data of the subtracter 1012 and stores the result in the frame memory 1004.

In this processing, depending on the amount of the data stored in the buffer circuit, the buffer circuit 1011 outputs a switching signal SP1 to the requantization circuit 1006 so as to switch the quantization step size of the re-quantization processing.

In this way, in said picture conversion apparatus 1001, the video data D0 is always transmitted at a constant rate.

Furthermore, in the picture conversion apparatus 1001 of the present embodiment, a motion field detection circuit 1014 decides the characteristic feature of the picture of each block. Then, based on this decision, the quantization step size is switched so that the high quality video data can be transmitted efficiently by making effective use of human visual sensitivity.

That is, as shown in Fig. 13, the motion field detection circuit 1014 successively inputs the detection result of the motion vector, and then, for each 3 x 3 block, performs the processing described by the following equation:

$$I_r = \sum_K |V_{i,j} \times V_K| + \sum_L \frac{|V_{i,j} \times V_L|}{\sqrt{2}} \quad \dots (1)$$

Where vector $V_{i,j}$ represents a motion vector V in the block $BL_{i,j}$ corresponding to a block which is to be processed by the re-quantization circuit 1006.

Subscripts K and L are given by

$$K = i, j - 1, i - 1, j, i + 1, j, i, j + 1 \quad (2)$$

$$L = i - 1, j - 1, i + 1, j - 1, i - 1, j + 1, i + 1, j + 1 \quad (3)$$

Thus, for each block BL, the motion field detection circuit 1014 detects the degree of disorder I_r of the motion vector V between the block BL and surrounding eight blocks BL.

That is, the vector products are obtained for the motion vector among the surrounding blocks BL, and then the sum of the absolute values of these vector products is obtained. If the motion vectors V show good similarity in their directions and magnitudes, then a small detection result I_r is obtained.

Therefore, when the calculation described by equation (1) is carried out for the surrounding eight blocks BL, if there is good similarity in the direction and the magnitude of the vectors V among the surrounding eight blocks BL, and then the motion field detection circuit 1014 gives a small detection result I_r .

That is, as shown in Fig. 14, in the case where the directions and magnitudes of the motion vectors are similar to each other among the surrounding blocks, it is possible to decide that the present block has a picture showing a motion similar to the motion in surrounding blocks. As a result, the small detection result I_r is contained.

On the other hand, when there is no similarity in the magnitude and direction of the motion vector V among the surrounding blocks as in Fig. 15, the present block can be considered to have a picture showing a motion different from those of surrounding blocks. Thus, the large detection result I_r is obtained.

In this way, the motion field detection circuit 1014 can detect the degree of disorder compared to the surrounding blocks based on the detection result I_r . Thus, the characteristic feature of the picture in each block compared to the surrounding blocks can be detected.

That is, if there is a similarity in motion between an picture in the present block and other pictures in surrounding blocks, then it is possible to decide that the entire picture moves uniformly as a whole. On the other hand, if there is difference in motion between a picture in the present block and other picture in surrounding blocks, it can be considered that the present block moves differently compared to other surrounding blocks.

In the case where the present block shows a different motion compared to the motion of other surrounding blocks, the present block has a characteristic feature that the degradation of the picture quality in the present block cannot be easily perceived.

Therefore, for the present area, the quantization step size is set to a large value, and then, as a result of this, a small quantization step size can be given for other areas. Thus, the degradation of the picture quality can be effectively avoided and efficient transmission of the video signal can be achieved.

Furthermore, in the process of detecting the degree of disorder I_r , the motion field detection circuit 1014 makes weighting is such a way that the vector products are weighted by the factor of $2^{-1/2}$ for the adjacent blocks existing apart in oblique the directions to the right and the left. Thus, the detection of the degree of disorder is carried out taking into account the effect of the distance between the control block BL_{ij} and other blocks.

In this way, the degree of the disorder I_r is detected taking into account the distance from the central block BL_{ij} , the detection of the disorder can become more realistic corresponding to the actual motion of the picture.

Therefore, if the quantization step size is switched based on the detection result of disorder, then higher efficiency can be achieved in transmission of higher quality of video signal by making more effective use of human visual sensitivity.

That is, the motion field detection circuit 1014 outputs the switching signal SP2 to the re-quantization circuit 1006 based on the detection result of the disorder I_r .

As shown in Fig. 16, by switching the table T1 to TN, the re-quantization circuit 1006 switches the quantization step size from Q_1 to Q_n according to the switching signal SP1 and SP2 so that the quantization step size $Q_1 - Q_n$ can be switched depending on the nature of a picture and on the amount of the data stored in the buffer memory 1011.

As shown in Fig. 17, a demodulation circuit 1701 gives the transmission data D0 to a decoding circuit 1703 via the buffer memory 1702. The decoding circuit 1703 performs the inverse processing to the processing performed by the variable length coding circuit 1010 so as to reproduce the output data of the re-quantization circuit 1006.

An inverse re-quantization circuit 1704 performs the inverse processing to the processing performed by the re-quantization circuit 1006 according to the switching signal SP1 and SP2 transmitted together with the transmission data D0. Thus, the input data of the re-quantization circuit 106 is reproduced.

A selection circuit 1705 switches the connection according to the selection signal SEL so that when the intraframe coding processing is selected, the output data of the inverse re-quantization circuit 1704 is directly output.

Thus, the demodulation circuit 1701 can reproduce the intraframe-coded video data D01.

A frame memory 1706 stores the output data of the selection circuit 1705. Thus, one-frame-previous video data is stored as in the video signal transmission apparatus 1.

On the other hand, when the interframe coding processing is selected, the selection circuit 1705 switches the connection so that the output data of an adder 1707 is selectively output.

The adder 1707 inputs the output data of the frame memory 1706 via a motion compensation circuit 1708, then adds the motion compensation video data with the video data coming from the transmission object. Then, the added result is output.

Thus, when the interframe coding processing is selected, the adder 1707 reproduces the input data of the adder 1012 so that the demodulation circuit 1701 reproduces the video signal transmitted via the selection circuit 1705.

(4-2) Operation of the Embodiment

In the above construction, the video signal is converted in a digital signal, and then applied to a motion detection circuit 1003 by a unit data of a block BL containing 8×8 pixels. Then, the motion detection circuit 1003 detects motion vector V using the block matching method.

This motion vector is applied to a motion compensation circuit 1013, by which the motion compensation is carried out for the one-frame-previous video data.

The motion compensated video data is applied to a subtracter 1012, by which the motion compensated video data is subtracted from the successively input video data to produce the residual difference data.

When the interframe coding processing is selected, a re-quantization circuit 1006 performs discrete cosine

transform for the residual difference data, then re-quantizes the resultant converted data. The resultant quantized data is output.

On the other hand, when the intraframe coding processing is selected, the re-quantization circuit 1006 performs discrete cosine transform directly for the successively input video data instead of the residual difference data, and then re-quantizes and outputs it.

In the above processing, a motion field detection circuit 1014 detects the degree of disorder of the motion vector V so that the quantization step size is switched according to the above detection result SP2.

When there is no similarity in the direction and the magnitude between the motion vector V in the present block and those in the surrounding blocks, the quantization step size is set large. As a result of this, it becomes possible to size is set large. As a result of this, it becomes possible to set the quantization step size small for other area.

In this way, the quantization step size is set large for the area where the degradation of the picture quality cannot be easily perceived. Thus, as a result of this, it becomes possible to set the quantization step size small for other area. Hence, high efficient transmission of high quality video signal can be achieved by making effective use of human visual sensitivity.

(4-3) Other Embodiments

In the above embodiment, the degree of disorder of the motion vector is detected within the area of 3 x 3 blocks. However, the present invention is not limited to that. The area where the detection is carried out can be set to any size as required.

Furthermore, in the above embodiment, an area of 40 x 24 pixels is set as a motion vector detection area for a block of 8 x 8 pixels. However, the present invention is not limited to that. The size of the block and the size of the motion vector detection area can be freely set as required.

Furthermore, in the above embodiment, the interframe coding processing and intraframe coding processing is switched for one unit block. However, the present invention is not limited to that. For example, the interframe coding processing and the intraframe coding processing can be switched for one frame. Moreover, other similar modifications are also possible.

While there has been described in connection with the preferred embodiments of the invention, it will be obvious to those skilled in the art that various changes and modifications as fall within the true spirit and scope of the invention.

Claims

1. A target object following apparatus comprising:
 - a picture data blocking circuit for dividing picture data sequentially input into blocks per unit of predetermined pixels in horizontal and vertical directions;
 - a detection area setting circuit for setting a tracking vector detection area including predetermined numbers of blocks in horizontal and vertical directions;
 - a motion detection circuit for detecting a motion vector for said respective blocks in said tracking vector detection area; and
 - a tracking vector detection circuit for detecting a distribution of said motion vector in said tracking vector detection area on the basis of the detection result of said motion detection circuit, and then detecting a moving direction and moving quantity of a portion of a picture in said tracking vector detection area on the basis of said distribution detection result.
2. The target object following apparatus according to claim 1, further comprising:
 - an imaging means for outputting a picture data obtained to image a predetermined target object into said target object following apparatus;
 - an imaging area changing means for changing an imaging area with said imaging means on the basis of a predetermined control signal; and
 - a detection area correction circuit for moving said tracking vector detection area with tracking to a motion of said picture in said tracking vector detection area, on the basis of the detection result of said tracking vector detection circuit within said target object following apparatus.
3. A target object following apparatus comprising:
 - a spatial edge detection circuit for outputting a spatial edge detection result for a motion picture

sequentially input;

a time edge detection circuit for outputting a time edge detection result for said motion picture; and
an edge judging means for outputting a motion edge detection result, on the basis of said spatial

edge detection result and said time edge detection result: wherein

5 said spatial edge detection circuit comprises;

a spatial power detection circuit for outputting a spatial power detection result detecting powers of
high frequency components per an unit of predetermined blocks of said motion picture; and

first comparison circuit for obtaining a comparison result of comparing said spatial power detection
result with first comparison reference, and then outputting said spatial edge detection result, when said
10 spatial power detection result is larger than said first comparison reference: wherein

said time edge detection circuit comprises;

a time power detection circuit for obtaining a difference data between the present frame picture
data and the previous frame picture data of said motion picture, and then detecting a power of said dif-
ference data to output a time power detection result per an unit of predetermined blocks; and

15 second comparison circuit for obtaining a comparison result of comparing said time power detection
result with second comparison reference, and then, when said time power detection result is larger than
said second comparison reference, outputting said time edge detection result: wherein

said edge judging circuit judges as an edge having a motion, when both of said spatial edge de-
tection result and said time edge detection result are obtained, and then outputting said motion edge de-
20 tection result.

4. The target object following apparatus according to claim 3, further comprising:

a picture data blocking circuit for dividing picture data sequentially input into first and second blocks
per unit of predetermined pixels in a horizontal and a vertical directions;

25 a detection area setting circuit for setting a tracking vector detection area including predetermined
numbers of said blocks in a horizontal and a vertical directions;

a motion detection circuit for detecting a motion vector for said second respective blocks within said
tracking vector detection area; and

30 a tracking vector detection circuit for detecting a distribution of said motion vector within said track-
ing vector detection area on the basis of the detection results of said spatial edge detection circuit and
said time edge detection circuit in said target object following apparatus, and then detecting a moving di-
rection and a moving quality of said picture within said tracking vector detection area on the basis of said
distribution detection result.

35 5. The target object following apparatus according to claim 4, further comprising:

an imaging means for imaging predetermined target object, and then outputting picture data into
said target object following apparatus;

an imaging area changing means for changing said imaging area of said imaging means on the
basis of predetermined control signal; and

40 a detection area correction circuit for tracking the movement of said picture within tracking vector
detection area on the basis of the detection result of said tracking vector detection circuit within said target
object following apparatus.

6. A picture conversion apparatus comprising:

45 a motion detection circuit for detecting a motion vector of respective blocks, after dividing picture
data input sequentially;

a picture data processing circuit for motion compensating said picture data on the basis of said mo-
tion vector, and then outputting residual data;

50 a motion field detection circuit for detecting disorder of said motion vector for a block of said residual
data and surrounding block adjacent said block;

a quantizing circuit for quantizing said residual data, and then outputting quantized data; and

a buffer circuit for storing said quantized data, and then outputting said stored quantized data;

wherein

said quantizing circuit for changing a quantizing step size in large way, when said vector is disorder,
on the basis of said motion field detection circuit.

55

7. A picture processing apparatus comprising:

said target object following apparatus according to claim 1, 2, 3, 4 or 5; and

said picture conversion apparatus according to claim 6.

8. The target object following apparatus according to claim 1, 4 or 6, wherein:
said motion vector is computed by means of block matching method.

5

9. The target object following apparatus according to claim 1 or 4, wherein:
in said tracking vector detection circuit,
data represented by numbers of said motion vectors in stored into tables which are provided for
respective moving directions and moving quantities within said tracking vector detection area,
10 the numbers of distributing of said table decide said moving direction and said moving quantity to
the maximum motion vector.

10. The target object following apparatus according to claim 4, wherein:
in said motion detection circuit, said motion vector is calculated only when said spatial edge data
said time edge exist.

15

11. The picture processing apparatus according to claim 6, wherein:
in said motion field detection circuit, said disorder of said motion vector is calculated such a manner
that it executes outer-product of said motion vector and the vector of surrounding block, and then executes
absolute summation of said outer-product results.

20

25

30

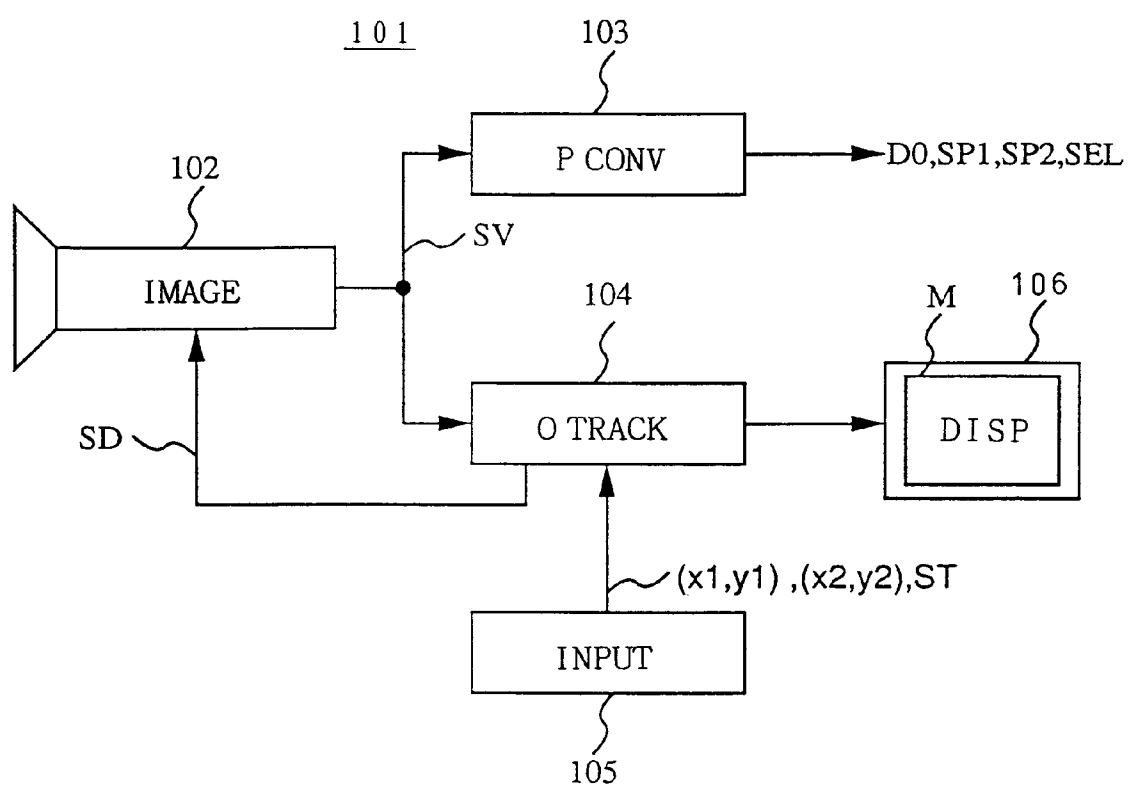
35

40

45

50

55

**FIG. 1**

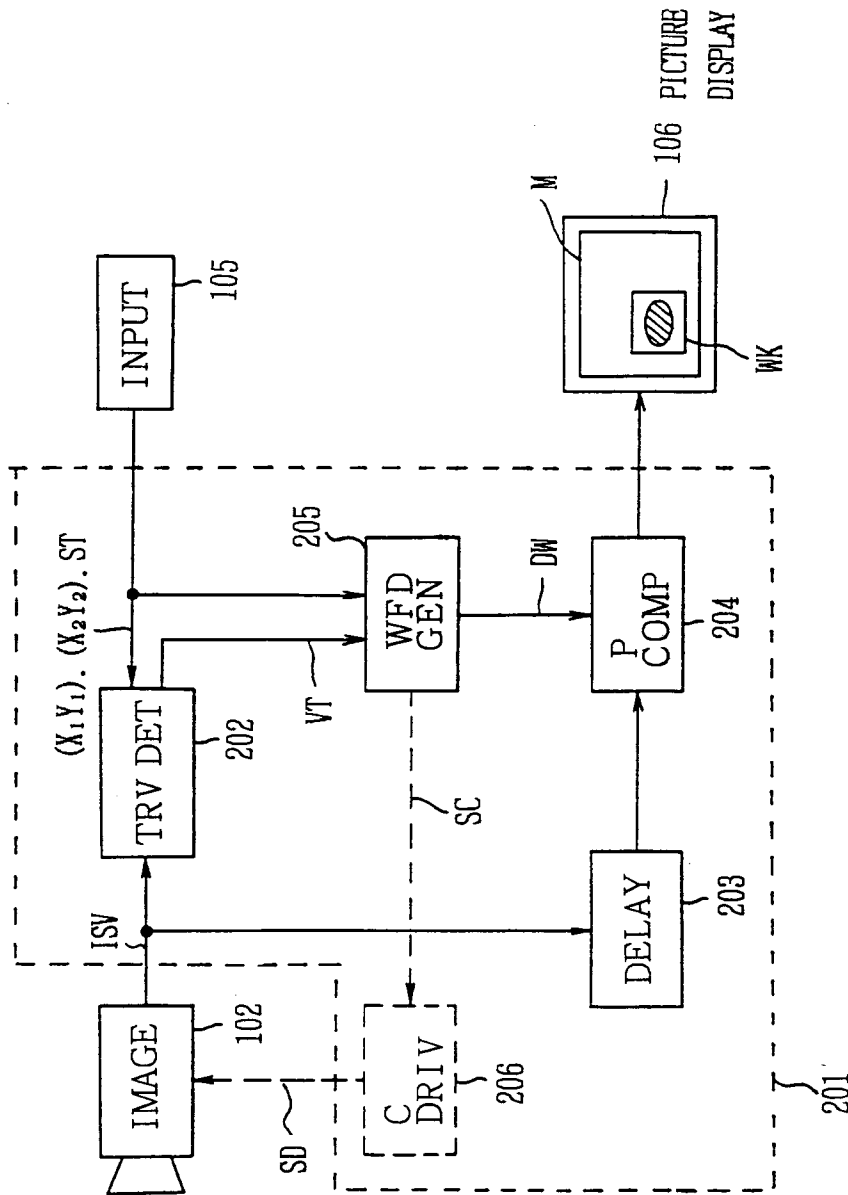


FIG. 2

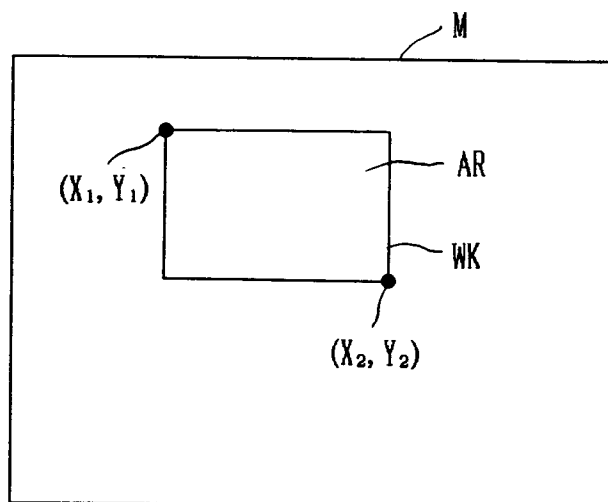


FIG. 3

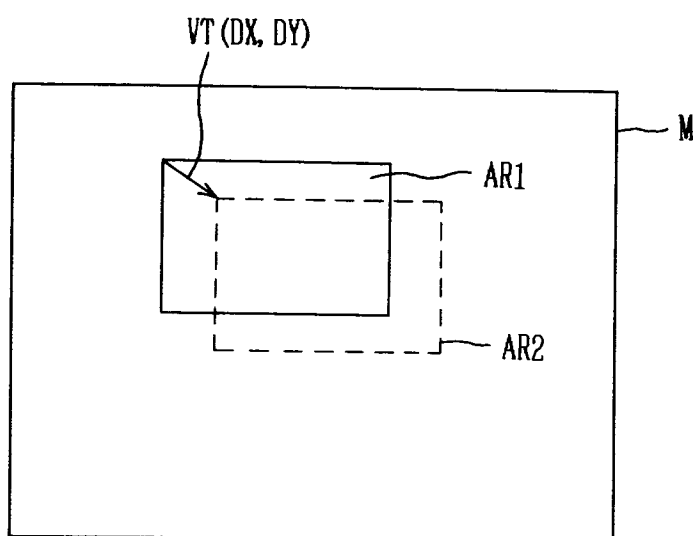


FIG. 4

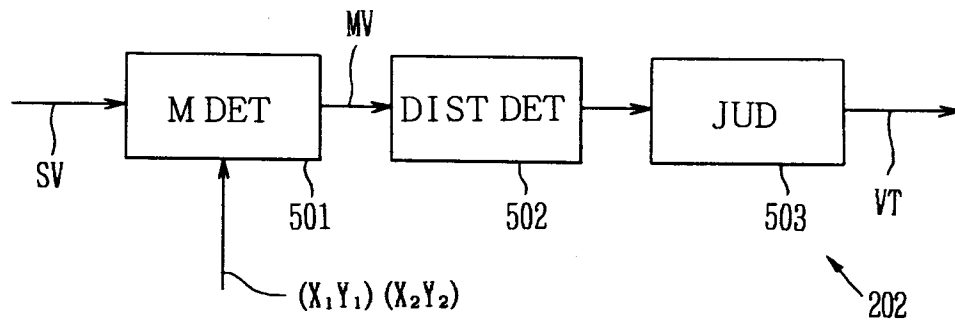


FIG. 5

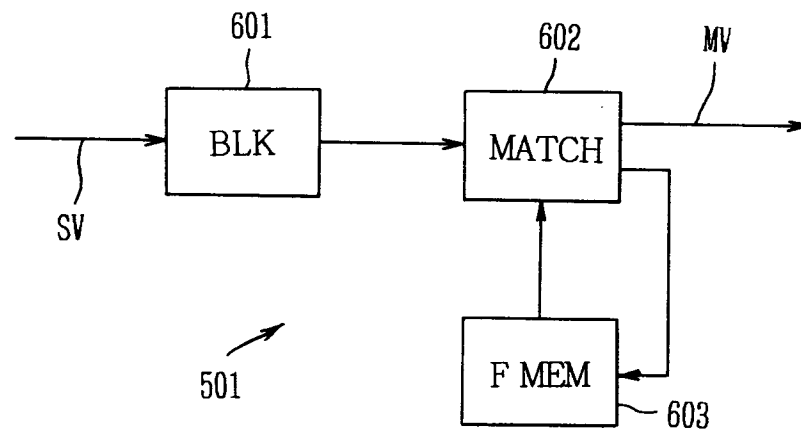


FIG. 6

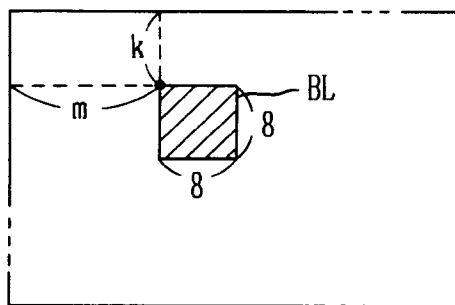


FIG. 7

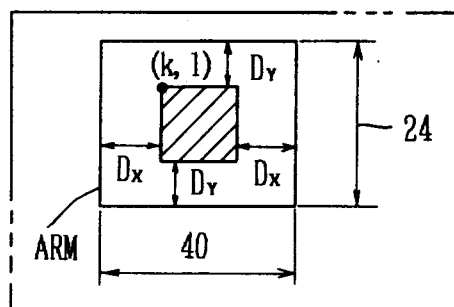


FIG. 8

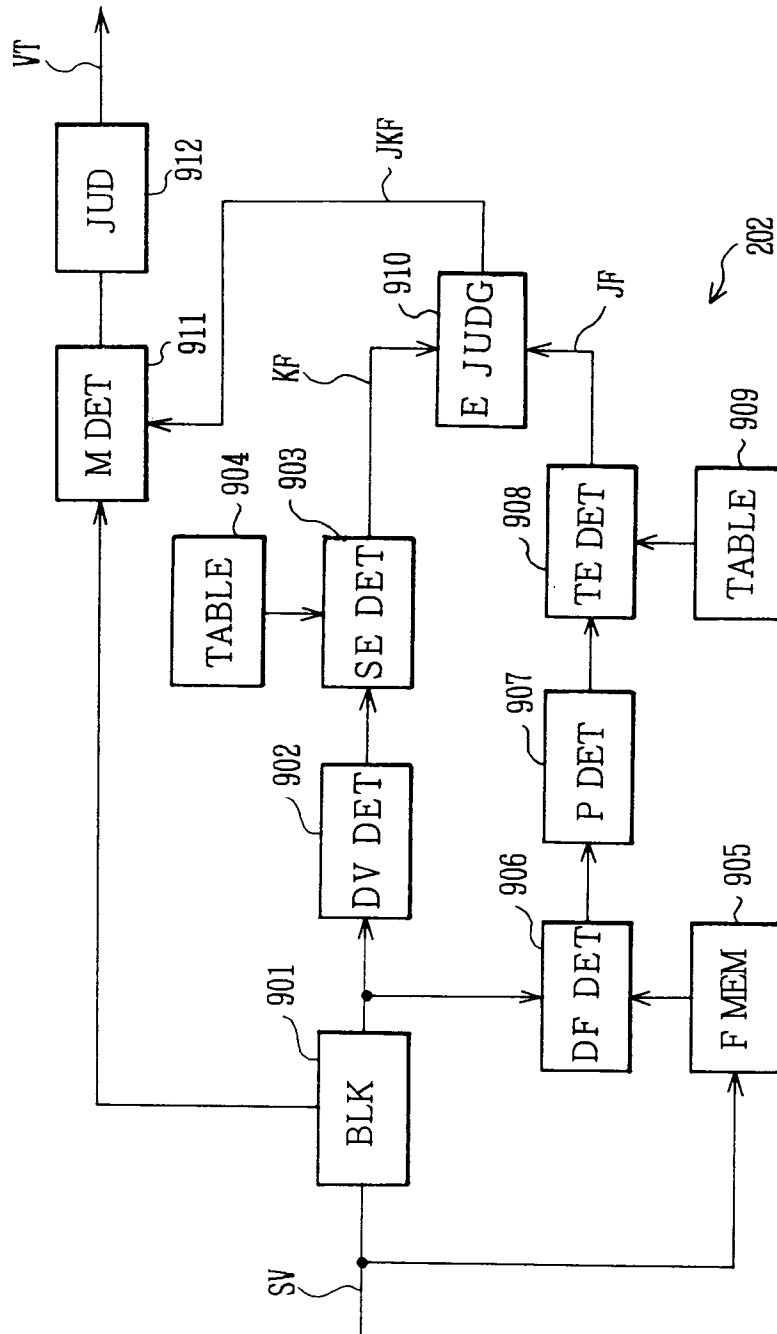


FIG. 9

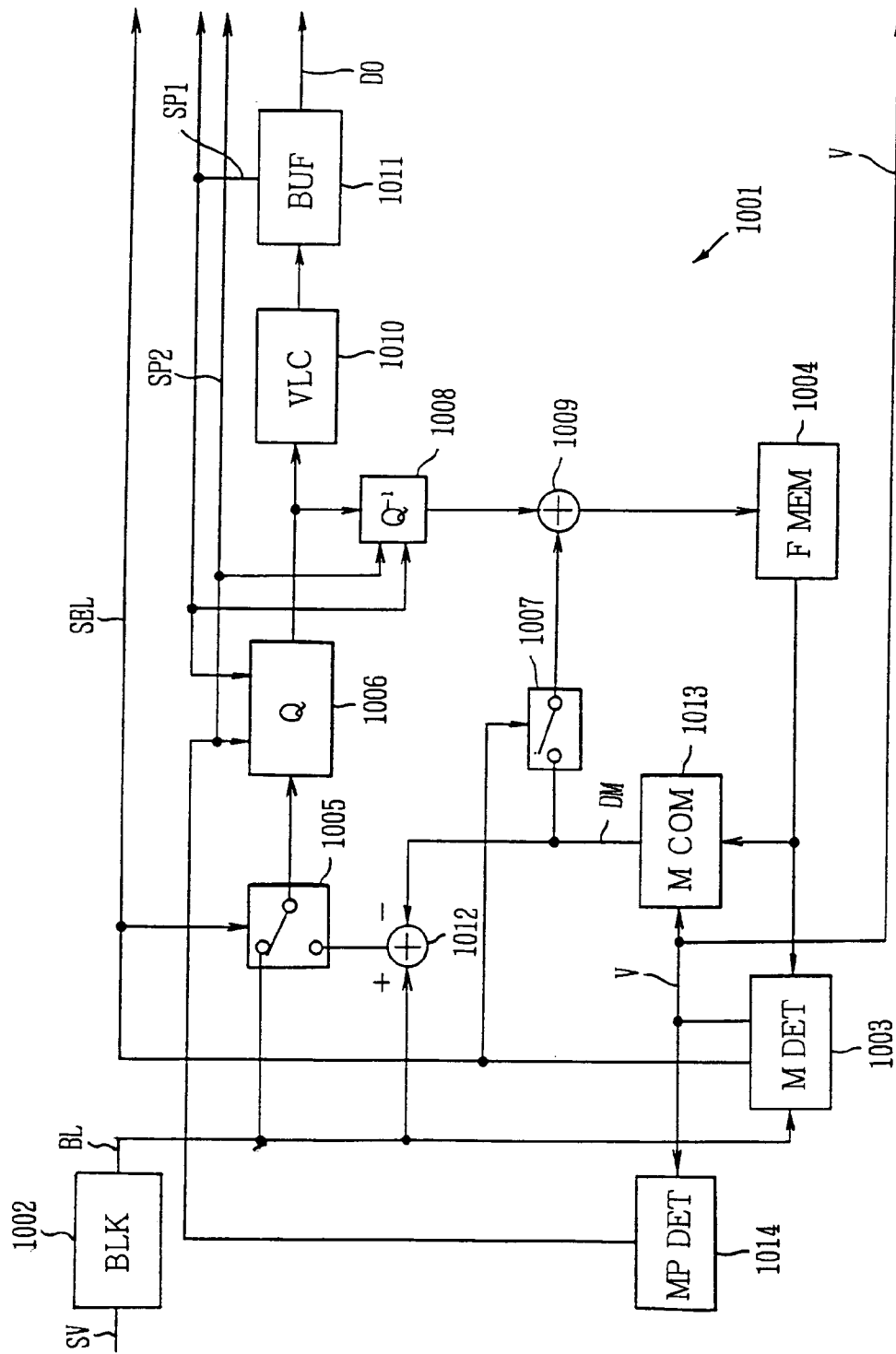


FIG. 10

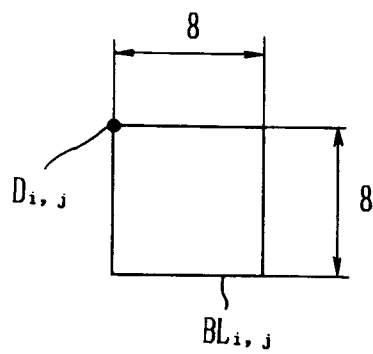


FIG. 11

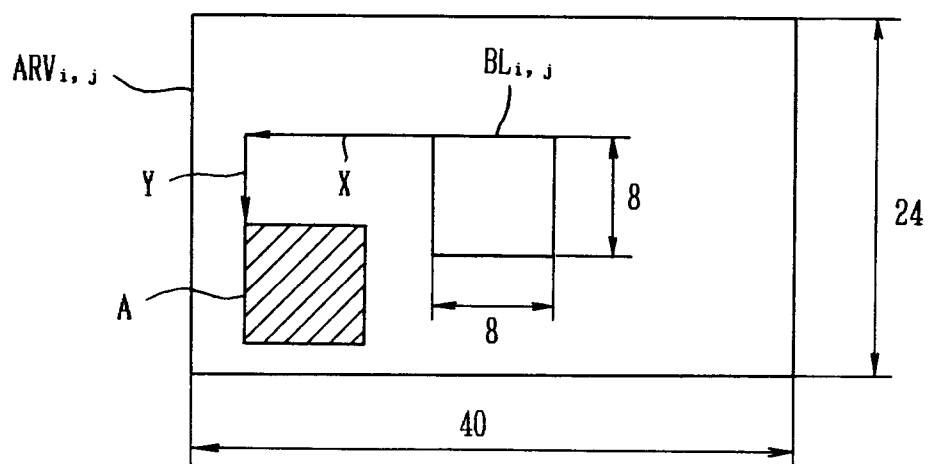


FIG. 12

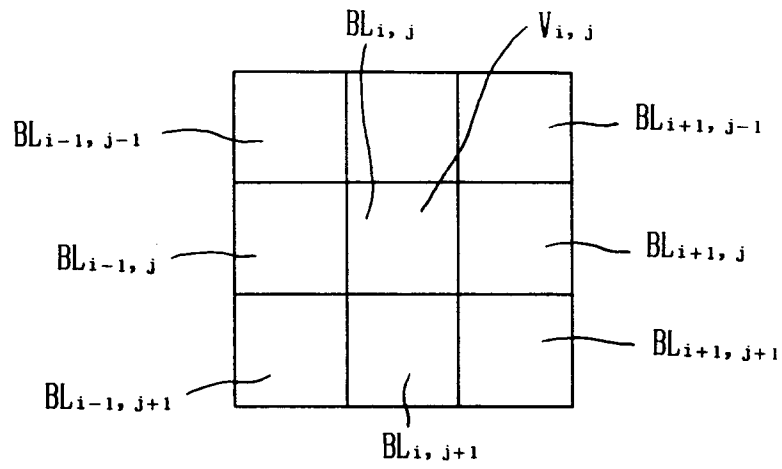


FIG. 13

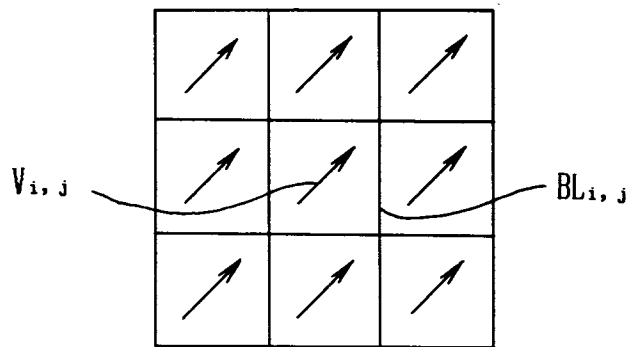


FIG. 14

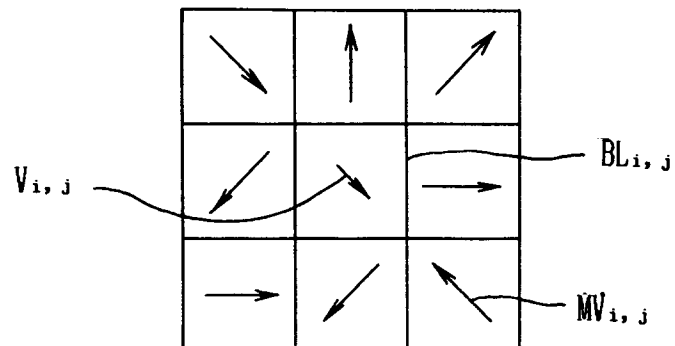


FIG. 15

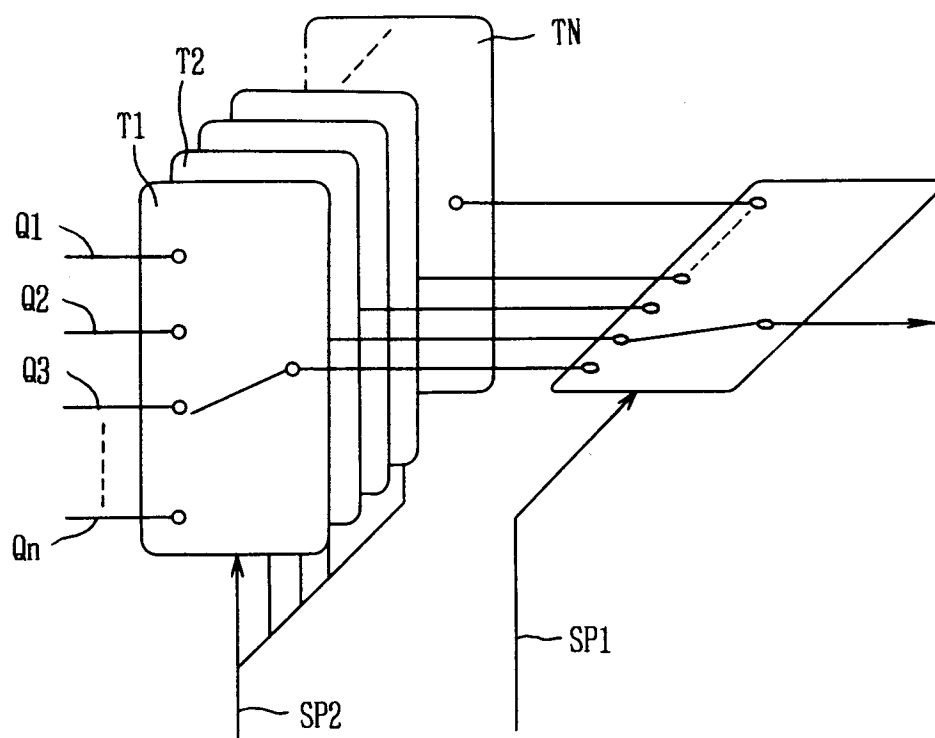


FIG. 16

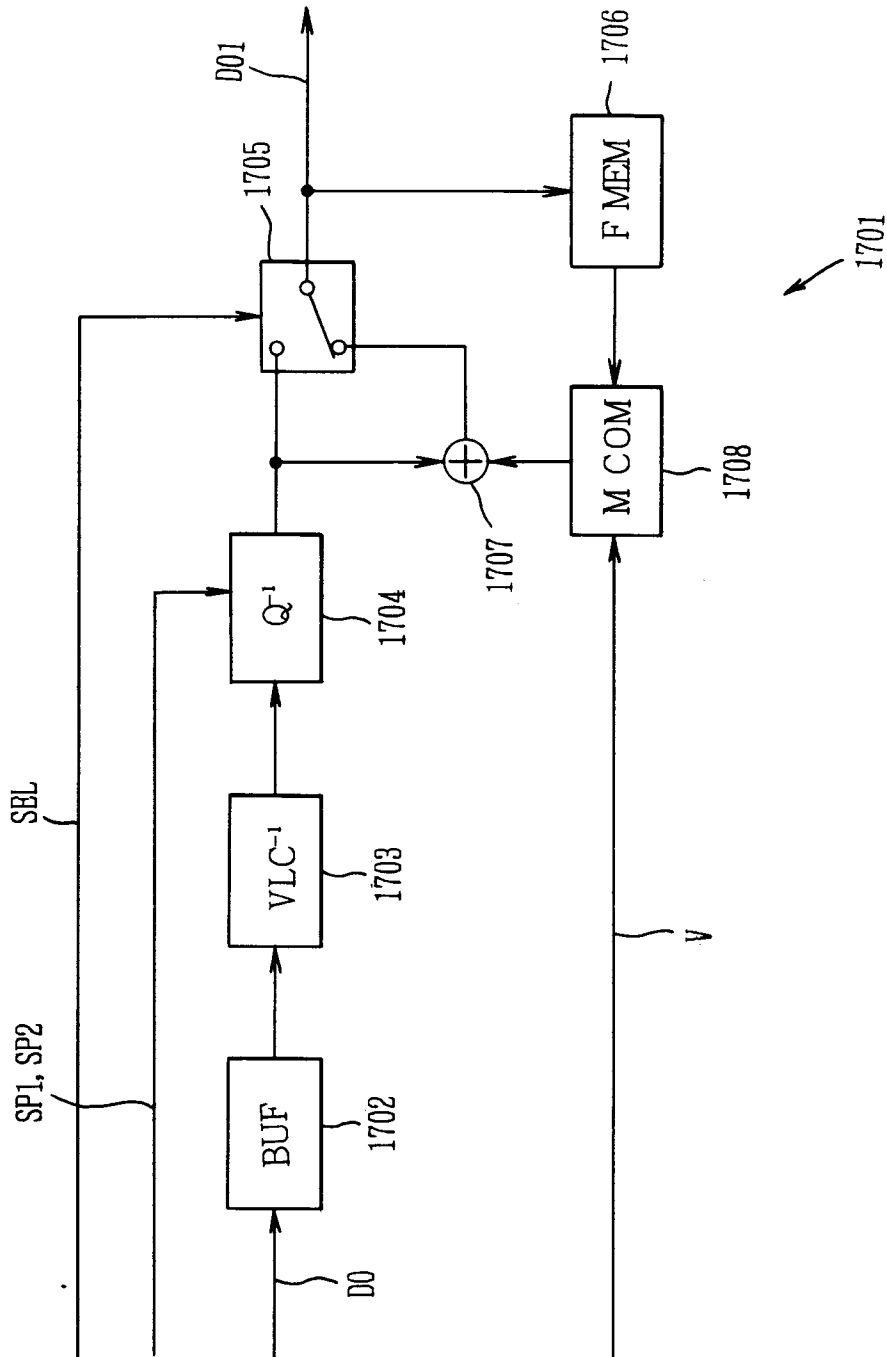


FIG. 17